Application Example 16
(Conditional second moment analysis with vectors)

PREDICTION OF DAILY TEMPERATURES USING SEVERAL PAST OBSERVATIONS

In Application Example 15, we have seen how one can use second-moment results for normally distributed variables to update uncertainty on a scalar quantity X_1 given a scalar observation X_2. For many applications, one needs to extend those results to the case when the predicted quantity and/or the observed quantity is a vector. We first review these extended results and then make an application to the prediction of temperature at different time lags.

Conditional Distribution Results for Jointly Normal Vectors

Consider two random vectors $\underline{X}_1$ and $\underline{X}_2$ with joint normal distribution, mean value vectors $\underline{m}_1$ and $\underline{m}_2$, and auto-covariance and cross-covariance matrices $\underline{\Sigma}_{11}$, $\underline{\Sigma}_{22}$ and $\underline{\Sigma}_{12} = \underline{\Sigma}_{21}^T$. These matrices are defined as $\underline{\Sigma}_{ij} = E[(\underline{X}_i - \underline{m}_i)(\underline{X}_j - \underline{m}_j)^T]$, i, j = 1, 2, where the superscript T denotes transposition. This means that the vector $\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix}$ has joint normal distribution

$$\underline{X} = \begin{bmatrix} \underline{X}_1 \\ \underline{X}_2 \end{bmatrix} \approx N\left(\begin{bmatrix} \underline{m}_1 \\ \underline{m}_2 \end{bmatrix}, \begin{bmatrix} \underline{\Sigma}_{11} & \underline{\Sigma}_{12} \\ \underline{\Sigma}_{21} & \underline{\Sigma}_{22} \end{bmatrix} \right) \qquad (1)$$

Now suppose that $\underline{X}_2$ is measured and found to be equal to $\underline{x}_2$. What is the conditional distribution of $(\underline{X}_1|\underline{X}_2 = \underline{x}_2)$? One can show that this conditional distribution is also normal, with mean value vector $\underline{m}_{1|2}$ and covariance matrix $\underline{\Sigma}_{1|2}$ given by

$$\begin{aligned} \underline{m}_{1|2} &= \underline{m}_1 + \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}(\underline{x}_2 - \underline{m}_2) \\ \underline{\Sigma}_{1|2} &= \underline{\Sigma}_{11} - \underline{\Sigma}_{12}\underline{\Sigma}_{22}^{-1}\underline{\Sigma}_{12}^T \end{aligned} \qquad (2)$$

In the special case when $\underline{X}_1$ and $\underline{X}_2$ are scalar quantities, $\underline{\Sigma}_{11} = \sigma_1^2$, $\underline{\Sigma}_{22} = \sigma_2^2$, and $\underline{\Sigma}_{12} = \underline{\Sigma}_{21} = \mathrm{Cov}[X_1,X_2] = \rho\sigma_1\sigma_2$ where ρ is the correlation coefficient between X_1 and X_2. Substitution into Eq. 2 gives

$$
\begin{aligned}
m_{1|2} &= m_1 + \rho\frac{\sigma_1}{\sigma_2}\left(x_2 - m_2\right) \\
\sigma_{1|2}^2 &= \sigma_1^2(1-\rho^2)
\end{aligned}
\qquad (3)
$$

which is the result in the scalar case (see Application Example 15). Like in the scalar case, in the case when the distributions of $\underline{X}_1$ and $\underline{X}_2$ are not normal the expression for $\underline{m}_{1|2}$ in Eq. 2 is not in general the conditional mean of $(\underline{X}_1|\underline{X}_2 = \underline{x}_2)$, but it always has the meaning of best linear unbiased estimator.

If one is not interested in the conditional covariances among the components of $\underline{X}_1$, then one may apply Eq. 2 separately to each component of $\underline{X}_1$. We do so in the application that follows.

Prediction of Daily Temperatures

Consider the sequence of maximum daily temperatures, $\{X_i, i = 0, \pm1, \pm2, ...\}$ (the X_i could equally well be daily stock prices, soil properties at different spatial locations, river discharges in different months, traffic volumes in different days, etc.). We observe X_i for $i = i_0$ (today), and several previous days i_0-1, i_0-2, ... We want to use these observations, with values x_{i_0}, x_{i_0-1}, x_{i_0-2}, ... to predict the future maximum daily temperatures X_{i_0+1}, X_{i_0+2}, ...

To be specific, suppose that we want to forecast daily temperatures in Boston, for the month of November. To do so, we need the mean values, variances and covariances of such maximum temperatures. From a historical record we obtain the following statistics:

- mean value (assumed to be the same for all days of the month): $m = 7°C$
- standard deviation (assumed to be the same for all days of the month): $\sigma = 5°C$

• correlation function ρ_{ij} (assumed to depend only on the time lag |i-j|):

|i-j|	ρ_{ij}	**|i-j|**	ρ_{ij}
0	1.00	6	0.16
1	0.95	7	0.08
2	0.81	8	0.04
3	0.63	9	0.02
4	0.44	≥ 10	0.00
5	0.28		

Using conditional second-moment analysis one can obtain several important results on temperature prediction. For example, one can predict temperature n days ahead, X_{i_0+n}, using only temperature today, X_{i_0}, or both temperature today and temperature yesterday. In the first case, calculations are based on Eq. 3. If one wants to make predictions based on temperature today and temperature yesterday, then one must use Eq. 2, where $\underline{X}_1 = X_{i_0+n}$ is a scalar and $\underline{X}_2 = [X_{i_0}, X_{i_0-1}]^T$ is a vector.

Numerical results are shown in Figures 1 and 2. For the case when temperature today is X_{i_0} = 15°C, Figure 1 shows $E[X_{i_0+n}|X_{i_0}$ = 15°C] - 7°C and $\{Var[X_{i_0+n}|X_{i_0}]\}^{1/2}$ as a function of the prediction time lag n. Notice that $E[X_{i_0+n}|X_{i_0}$ = 15°C] - 7°C is the amount by which the conditional mean (best predictor) deviates from the unconditional seasonal mean of 7°C and $\{Var[X_{i_0+n}|X_{i_0}]\}^{1/2}$ is the standard deviation of the prediction error. As one can see from Eq. 3, $E[X_{i_0+n}|X_{i_0}$ = 15°C] - 7°C is proportional to the correlation coefficient ρ_n. For n ≥ 10, $\rho_n = 0$ and the best predictor is the seasonal mean. This is why, for n ≥ 10, $E[X_{i_0+n}|X_{i_0}$ = 15°C] - 7°C = 0 and $\{Var[X_{i_0+n}|X_{i_0}]\}^{1/2}$ = 5°C.

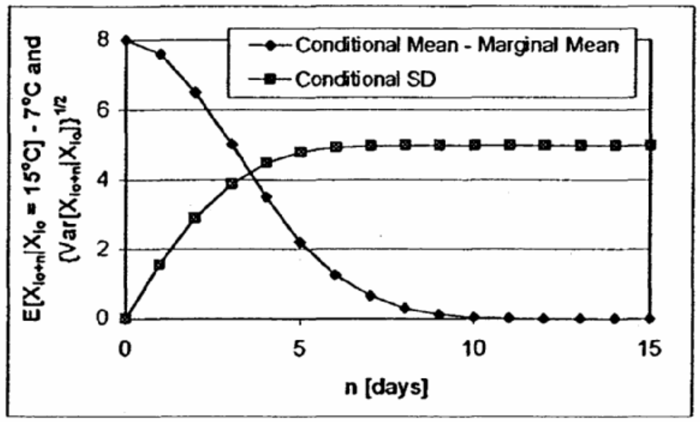

Figure 1: $\mathbf{E[X_{i_0+n}|X_{i_0} = 15°C] - 7°C}$ **and** $\mathbf{\{Var[X_{i_0+n}|X_{i_0}]\}^{1/2}}$ **as a function of prediction lag in days, n**

Figure 2 shows similar results when prediction is based on the observation of temperature today and yesterday. To exemplify, it was assumed that the observed temperature is 15°C for both days. While the trend of the conditional mean and conditional standard deviation are similar to those based only on temperature today, the values are not exactly the same. In particular, using information about temperature yesterday reduces the standard deviation of the prediction error (for example, the standard deviation for two-day lag prediction is about 3°C when one uses only temperature today and about 1.8°C when also temperature yesterday is used).

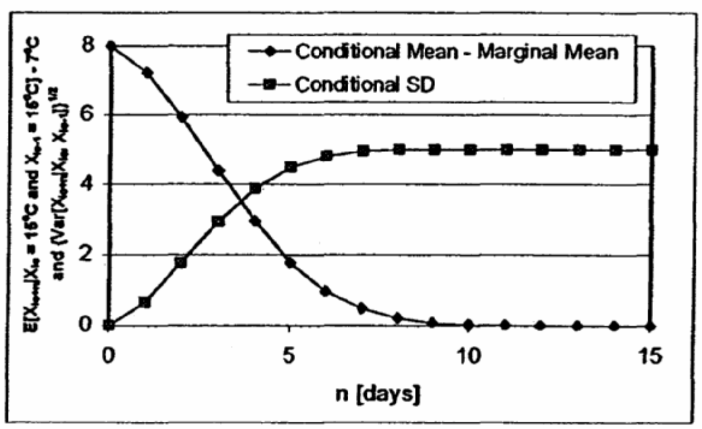

Figure 2: $\mathbf{E[X_{i_0+n}|X_{i_0} = 15°C \text{ and } X_{i_0-1} = 15°C] - 7°C}$ **and** $\mathbf{\{Var[X_{i_0+n}|X_{i_0}, X_{i_0-1}]\}^{1/2}}$ **as functions of prediction lead in days, n**

Problem 16.1

Retain all the parameters of the example above, except for the correlations, which are now as follows.

$\|i-j\|$	ρ_{ij}	$\|i-j\|$	ρ_{ij}
0	1.0	6	0.4
1	0.9	7	0.3
2	0.8	8	0.2
3	0.7	9	0.1
4	0.6	≥ 10	0.0
5	0.5		

Produce plots analogous to those in Figures 1 and 2. Compare your results with those in Figures 1 and 2, giving qualitative explanations for the differences.

[Note: You may find that, for the new correlation function, predictions using temperature today or temperature today and yesterday are very close. Such predictions are exactly the same if the temperature sequence has a property called Markovian dependence. For a Gaussian Markov sequence, the correlation function decays in an exponential manner, i.e. as $\rho_{ij} = e^{-|i-j|/c}$, where c is a positive constant. The linear correlation function you have used is not very different from exponential. This is why for that correlation function predictions are insensitive to considering temperature readings in previous days.]